The Metacognitive Monitoring and Study Decisions of Incremental Theorists

Laura Brumby

A report submitted as a partial requirement for the degree of Bachelor of Psychological Science with Honours at the University of Tasmania, 2016

Statements of Sources

I declare that this report is my own original work and that contributions of others have been duly acknowledged.

Signed: ______________________ Date: 13/10/2016

Acknowledgments

Thank you to my supervisor, Dr. Matthew Palmer, for all the support you have provided me throughout the year. I really appreciate all the time you have devoted to this project (in its many forms), and your assistance in calculating the statistics required for my calibration analyses. Thank you for all your guidance, patience and optimism, and ultimately, for making this process such a rewarding experience. Your enthusiasm for research is infectious, and it is a privilege to learn from you.

Thank you to Frances Parkes for your assistance in programming the learning task. The many hours of work you contributed, as well as all the troubleshooting, were very much appreciated.

Thank you to Nicole McCallum for recruiting and running participants at Flinders University.

To all the psychology staff in Launceston, thank you for the various ways you have supported me throughout the year. Especially, thank you Peter Tranent for your advice on an earlier version of the study, and Dr Christine Padgett for assisting with participant recruitment. Thank you to Dr Cynthia Honan for the encouragement and support you have shown me.

Additionally, thank you to Karen Palmer for your valued input on an earlier version of this study as well.

Thank you also to the psychology post-graduate candidates, Glenys Holt and Daniel Zuj. I have really valued the guidance and advice you have offered me this year, as well as your willingness to assist me where possible.

I would also like to acknowledge my peers; your friendship and support have been great assets.

Thank you to all the people who generously volunteered their time to participate in my research; without you, this study would not have been possible.

Thank you to my friends who have supported me along this journey in various ways as well.

And last, but certainly not least, thank you to my wonderful family, who unwaveringly support me and inspire me to do better. I truly appreciate your willingness to help me in any way possible. Thank you especially to Mum and Dad for the practical ways you have contributed to this project, such as proof-reading drafts and recruiting participants. Matthew and Brianna, thank you for always being on-call for assistance too. A special thanks also goes to Henrick for teaching me various computer skills, which have come in super-handy.

Table of Contents

List of Tables

List of Figures

The Metacognitive Monitoring and Study Decisions of Incremental Theorists

Laura Brumby

Word Count: 9005

Abstract

The present study investigated incremental theorists' metacognitive monitoring and study behaviours during a word-pair learning task. Sixty-five participants (38 female; aged 18-64 years, *M* = 33.92, *SD* = 14.40) studied English-Spanish word-pairs of differing difficulty (easy, moderate and difficult) and made judgments of learning (JOLs). Participants selected half of the word-pairs for restudy, before completing an initial cued-recall test to demonstrate the accuracy of their JOLs and study decisions. They then completed a restudy phase, incorporating a within-in subjects honour/dishonour manipulation, to assess the optimality of their study decisions. A final cued-recall test was administered, followed by Dweck's (1999) Theories of Intelligence Scale. Endorsing more incremental views of intelligence was marginally associated with poorer calibration for difficult word-pairs, ($r = .235$, $p = .062$), and was negatively associated with metacognitive accuracy for easy word-pairs during the study choice phase ($r = -.260$, $p = .036$). No further impairments were evident for the other difficulty-levels of these statistics, nor on over-confidence, under-confidence or resolution. Participants' study choices reflected a Discrepancy Reduction approach, and in general, did not appear to be optimal. Future research is required to further clarify these findings in reference to a larger entity theorist comparison group.

People’s experiences of the world are not only determined by the external stimuli they encounter, but also by the meaning-making systems they use to perceive information through (Bartlett, 1932). This helps to explain why different people can respond to the same events in very different ways, as the interpretations they make inform their subsequent behaviour (Dweck, 1999). In an educational setting, it is of interest to know why some students adopt behaviours which have positive effects on their learning outcomes, such as seeking out challenge and persisting in the face of set-backs (Dweck, 1999), whilst others demonstrate less-optimal approaches, such as challenge avoidance and “helpless” responses to failure (Diener & Dweck, 1980; Dweck, 1999, Robins & Pals, 2002). One account for this observation is that students hold differing beliefs regarding the nature of intelligence, and their understanding forms a perceptual framework through which they make sense of their own experiences, and subsequently choose to engage in particular behaviours (Dweck & Leggett, 1988; Dweck, 1999).

According to Dweck and colleagues’ social-cognitive model, people hold implicit beliefs regarding their own intelligence (Dweck & Leggett, 1988; Dweck, Chui & Hong, 1995; Dweck, 1999). Two opposing views of intelligence are said to exist: an ‘*entity*’ theory, and an ‘*incremental*’ theory. People holding an entity theory of intelligence (entity theorists) consider intelligence to be a fixed and stable internal trait, which ultimately can never be changed or improved. Incremental theorists however view intelligence as being a malleable attribute, which does have the capacity to be developed over time. Over the past three decades, research has demonstrated that the implicit beliefs people hold have implications not only for their self-concepts, but also the attributions they make, the behaviours they exhibit, and their educational outcomes (eg. Dweck & Leggett, 1988; Hong, Chui, Dweck, Lin & Wan, 1999; Robins & Pals, 2002; Claro, Paunesku & Dweck, 2016).

Differences between entity and incremental theorists

One of the primary differences established between entity and incremental theorists is the type of motivational goal they show a preference for (Dweck, 1999). Possessing an entity view of intelligence is associated with the pursuit of performance goals (Dweck & Leggett, 1988). Performance goals involve engaging in tasks in such a way that allows one to demonstrate their abilities, in order to illicit favourable impressions. Furthermore, pursuing performance goals involves avoiding situations or experiences that may reveal weaknesses or low ability. This is especially important to the entity theorist, who would consider one demonstration of inadequacy as indicative of an enduring internal flaw (Dweck, 1999). Consequently, entity theorists prefer easier tasks which carry minimal risk of them experiencing difficulty or failure (Dweck, 1999). Incremental theorists however are less concerned with displaying their intellect, and more interested in tasks which will allow them to develop their knowledge and abilities (Dweck & Leggett, 1988). They therefore prefer to pursue learning goals, characterised by challenging experiences which offer the opportunity to learn something new (Dweck, 1999). The influence of theories of intelligence on goal preference has not only been implied by correlational research, but also established through experimental research in which entity or incremental views of intelligence have been manipulated in students, and their subsequent task choice reflected the goal-orientation described above (for reviews, see Dweck and Leggett, 1988; Dweck, 1999).

Theories of intelligence also influence the meaning students derive from needing to exert effort in order to complete a task (Dweck, 1999). For example, students have been demonstrated as judging others' intelligence differentially, based on their own theory of intelligence (Hong et al., 1999, Study 3). Hong and colleagues manipulated college students to have either an entity or incremental view of intelligence, and then asked them to compare the intelligence of two students in a scenario; one student put high levels of effort into their study, whilst another put minimal effort into studying. The majority of entity theorists (77%) claimed the student who put minimal effort into their study was the most intelligent. By contrast, incremental theorists were just as likely to consider either student as being the most intelligent (50% for each student), suggesting that they did not systematically consider effort to be an indicator of low ability, as the entity theorists did. These findings are congruent with the assertion that entity theorists consider the experience of effort exertion to indicate they have reached the limits of their ability, whereas incremental theorists consider effort to be a signal that they are developing their abilities (Dweck & Leggett, 1988; Dweck, 1999).

Implications of incremental orientation

Adaptive responses to failure. The interpretive frameworks that incremental theorists use to understand their performance on tasks sets them up to respond differently to failure than entity theorists. Endorsement of an entity theory of intelligence is associated with helpless responses, whilst incremental theorists take a mastery-oriented approach, perceiving achievement as reliant on effort (Robins & Pals, 2002; Dweck, 1999). Research has demonstrated that when given difficult tasks, children with helpless orientations exhibit negative affect and self-cognitions, and a decline in performance over trials (Diener & Dweck, 1980). In particular, Diener and Dweck found these children tended to spontaneously attribute their failure to their own internal inadequacies, such as low intelligence. Helpless children also diverted their attention away from the challenging task. Mastery-oriented children however did not display this pattern; rather, the challenging tasks motivated them to try harder, and ultimately thrive, as the majority of children (80%) maintained or improved their performance levels. Importantly, they did not disengage with the tasks or attempt to avoid the challenges. The finding that students operating under an incremental framework respond to failure more adaptively than their entity-oriented counterparts has been replicated with university students. For example, when comparing students who reported obtaining either poor or adequate grades, Hong and colleagues (1999, Study 2) found entity theorists with poor grades were not more likely to express interest in a remedial course, despite having a clear need for intervention. Incremental theorists reporting low grades did express a greater interest in the remedial course than both incremental theorists with high grades, and entity theorists in general, suggesting they respond more adaptively to failure than their entity theorist counter-parts.

Superior educational outcomes. Given that students who endorse incremental views of intelligence are considered to respond more adaptively in the short-term to experiences of challenge and failure than their entity theorist counter-parts, it is of little surprise that the literature consistently finds superior educational outcomes for incremental theorists (eg. Henderson & Dweck, 1990; Good, Aronson & Inzlicht, 2003; Blackwell, Trzesniewski & Dweck, 2007). For example, Henderson and Dweck (1990) followed students through their transition into junior high school (from 6th to 7th grade), a period characterised by posing many new and difficult challenges. Their research revealed entity theorists entering the 7th grade with either low or high levels of achievement obtained equal or worse marks (respectively) at the end of the year. Incremental theorists however either improved or maintained their marks, depending on whether they had entered the 7th grade with low or high marks, respectively. Thus, having an incremental theory of intelligence seemed to offer protection from the decline in achievement that might otherwise occur (Henderson & Dweck, 1990). Similar results have been observed by Blackwell and colleagues (2007, Study 2), who demonstrated that teaching 7th grade students to adopt an incremental view of intelligence protected them from the downward trajectory in math grades experienced by their peers in a control group.

Impressively, the academic benefits of having an incremental theory of intelligence are particularly pronounced for students who would typically be considered to be disadvantaged, or ‘at risk’ of receiving poor grades (Aronson, Fried & Good, 2002; Claro et al. , 2016). For example, Aronson and colleagues (2002) demonstrated that encouraging African American college students (who generally receive lower grades than their White classmates) to promote an incremental view of intelligence resulted in them being awarded higher end of year marks than the African American participants in both active and inactive control groups. Similarly, in a study of all 10th grade students in government schools in Chile, Claro and colleagues (2016) found that endorsing incremental views of intelligence (as opposed to entity-oriented views) was associated with higher academic performance across every level of socio-economic status. Moreover, endorsing incremental views appeared to offer low-income students some protection against the relationship between low-income disadvantage and low academic achievement.

Impaired metacognitive monitoring? Despite the aforementioned advantages of being an incremental theorist, there is some evidence to suggest that a short-term drawback may exist (Miele & Molden, 2010; Miele, Finn & Molden, 2011). Specifically, it seems that incremental theorists may experience distortions to their metacognitive monitoring (Miele, Finn & Molden, 2011). Metacognitive monitoring describes the process by which people are aware of their own cognitive experiences (Flavell, 1979). One way of measuring this is by asking people to make confidence judgements, such as *judgments of learning* (JOLs), which indicate how likely they think it is that they will be able to recall currently available information when tested in the future (Nelson & Dunlosky, 1991). High JOLs indicate that one is relatively confident of being able to produce the information on a future test, whilst low JOLs reflect a belief of being unlikely to recall the information. Confidence ratings can provide useful information about the likelihood of future recall (e.g. Koriat, 2008; Kornell & Bjork, 2008). It is important to acknowledge however, that confidence judgements are prone to bias from a variety of sources (Finn & Tauber, 2015). One such source of influence is encoding fluency, which refers to the feelings of ease or difficulty experienced when trying to learn new information. The common finding that feelings of ease at encoding result in high JOLs, whilst feelings of difficulty at encoding result in low JOLs, suggest that people tend to follow an *'easily-learned, easily-remembered'* (ELER) heuristic (Koriat, 2008; Koriat & Ma'ayan, 2005). When information varies objectively in difficulty, whereby some material will require more effort to learn than other material, this heuristic is likely to lead to accurate judgements about future recall (Koriat, 2008).

Given that entity theorists consider high levels of effort exertion to be indicative of low abilities, it is of little surprise that they tend to follow the ELER heuristic (Miele & Molden, 2010; Miele, Finn & Molden, 2011). In a study by Miele and colleagues (2011, Study 1), participants were given Indonesian-English language word-pairs to learn (each of which varied in difficulty, between easy, moderately difficult, or difficult to learn), and were asked to make a prediction (JOL) regarding their likelihood of recalling a target (English) word when presented with the cue (Indonesian) word from each word-pair. Participants endorsing an entity theory of intelligence were well calibrated in their JOLs, revealing that at each level of item difficulty, their JOLs reflected their actual performance. As item difficulty was negatively related with encoding fluency, this suggests that entity theorists followed the ELER heuristic, which was in this case, an informative approach. Incremental theorists however did not appear to use the ELER heuristic. Rather, their judgements reflected an inversed heuristic (*highly-engaging, easily-remembered, HEER*), whereby they were under-confident on the easy items, and over-confident on the difficult items (Miele et al., 2011). Incremental theorists therefore considered the materials requiring high levels of effort to encode (which were objectively difficult) as well-learned, whilst materials requiring lower effort exertion (which were objectively easier) as less-well learned. This pattern of results has also been observed in research in which effort cues were not derived from differing levels of objective difficulty between stimuli (Miele & Molden, 2010, Studies 4 & 5). Thus, it seems that incremental theorists' positive interpretation of effort may make them vulnerable to misguided metacognitive monitoring.

As metacognitive monitoring influences decision making processes involved in study behaviours (Nelson & Narens, 1990; Schwartz & Perfect, 2002), it is conceivable that incremental theorists' ability to engage in optimal study may be compromised. In order to assess whether this is the case, one must first establish what optimal study actually is. The literature has been predominantly divided between two accounts of optimal study: a Discrepancy Reduction (DR, eg. Nelson & Leonesio, 1988; Dunlosky & Thiede, 1998; Nelson & Narens, 1990) model, and a

Region of Proximal Learning (PRL, Metcalfe & Kornell, 2005; Kornell & Metcalfe, 2006) account.

The DR models suggest that people choose to allocate their study time preferentially towards the items that are considered to be most difficult (Nelson & Leonesio, 1988). Put another way, the focus of study is to reduce the *greatest* discrepancies between actual and desired knowledge. In three experiments of self-paced study, Nelson and Leonesio (1988) observed a negative correlation between confidence ratings and study time, indicating that students were allocating the most time to the items they reported as being the most difficult. Due to multiple replications in the literature (for a review, see Son & Metcalfe, 2000), and its intuitive appeal, the DR model became the dominant account of study allocation. The apparent popularity of the DR model however, does not necessarily indicate that it is an optimal approach to take. Nelson and Leonesio's (1988) own data demonstrated a negative correlation between study time and performance, suggesting this strategy did not yield optimal results.

The DR model has also been critiqued for being based upon research in which students were allowed unlimited study time (Kornell & Metcalfe, 2006). Kornell and Metcalfe argued that students are typically required to study under time constraints, whereby allocating time for studying one item comes at the expense of studying another. When this is the case, and mastery is therefore unlikely, they claim that a different approach to study is likely to occur. The RPL framework suggests that in such conditions, students should focus their study on the easiest items that they have not yet learned (Metcalfe & Kornell, 2005; Kornell & Metcalfe, 2006). Indeed, when participants are given limited study time, research suggests they focus on the unlearned items that they have given the highest confidence ratings to (eg. Son

& Metcalfe, 2000, Studies 1 & 3). Similarly, in studies where participants are forced to select a limited number of items for restudy, the majority select items receiving higher confidence ratings (Metcalfe & Kornell, 2005, Study 6; Kornell & Metcalfe, 2006, Studies 3a & 3b). Moreover, final test performance was greater for participants who followed this RPL approach than for those who did not (Kornell & Metcalfe, 2006, Studies 3a & 3b). It is important to note however, that these studies advocating the optimality of the RPL framework only allowed participants to choose from items that they had demonstrated they did not know on a preliminary test. This model therefore assumes that students are able to accurately discern between items that are known and unknown, which is consistent with research findings (e.g. Masur, McIntyre & Flavell, 1973).

Given the research suggesting that incremental theorists have compromised metacognitive monitoring, particularly in judging items that have and have not been learned (Miele et al., 2011, Study 1), it is conceivable that they may have difficulty following a RPL framework as successfully as their entity theorist counterparts. The current study therefore aims to investigate this possibility.

The present study.

The present study had three primary aims (outlined in the following paragraphs). The first was to replicate Miele and colleagues' (Miele & Molden, 2010; Miele et al., 2011) finding of compromised metacognitive monitoring for incremental theorists, compared to entity theorists, when learning items of differing objective difficulty. Assuming incremental theorists made inaccurate initial confidence judgments, the next aim was to investigate what impact this would have on their study behaviour. Specifically, it was of interest to know whether incremental theorists would have greater difficulty than entity theorists in selecting unknown items for restudy. Finally, the present study sought to investigate which study approach (DR or RPL) would lead to superior performance on a final cued-recall test.

It was hypothesised that participants endorsing incremental views of intelligence would make JOLs which would not be predictive of their actual knowledge demonstrated on an initial test. Based on the HEER heuristic (Miele et al., 2011), it was anticipated that incremental theorists would under-estimate their learning of objectively easier (and less engaging) stimuli, and over-estimate their learning of difficult (and highly engaging) stimuli. As entity theorists have been shown to follow the ELER heuristic (Miele et al., 2011), which provides valid learning cues when items differ in objective difficulty (Koriat, 2008), it was anticipated that their JOL ratings would better reflect their actual performance.

In regards to the consequence of compromised metacognitive monitoring on study behaviours, two possible outcomes were anticipated. Firstly, if incremental theorists had difficulty assessing the extent to which they had learned items, then they may have difficulty appropriately selecting items for restudy. In this case, incremental theorists would be expected to show poorer resolution than entity theorists; that is, they would be less able to accurately distinguish between known and unknown items. Alternatively, any differences that exist between incremental and entity theorists' metacognitive monitoring immediately after an item's presentation may disappear by the time they make their study decisions. Indeed, research has shown that confidence judgements tend to be less biased when they are taken after a delay in time, rather than during study (*'delayed-JOL-effect'*, Nelson & Dunlosky, 1991). As participants made their study choices after all items were presented, these decisions were effectively delayed JOLs, thus any impairment in incremental theorists' monitoring might be cancelled out.

Finally, as the design of the current study limited participants' study time, it was anticipated that following a RPL framework (rather than a DR model) would be the most popular, and optimal, approach. Therefore, it was hypothesised that the mean JOLs made for items that were later selected for restudy would be greater than the mean JOLs of items not selected for restudy. Moreover, it was expected that there would be a positive relationship between the mean JOLs for items chosen for restudy and final test performance, whereby greater average JOL ratings for the items restudied were anticipated to be associated with higher recall performance on the final test.

Method

Participants

Sixty-eight participants were recruited for the present study from the communities of the University of Tasmania and Flinders University. Due to non-compliance with experimental instructions, three participants were excluded. The final sample comprised 65 participants (38 females), ranging in age from 18-64 years (M_{age} = 33.92 years, *SD* = 14.40). Participants were required to report being able to fluently read and write in English, and to not be fluent in Spanish. Participants were either reimbursed $20 for their time, or awarded course credit.

Materials

The current study was presented to all participants on desktop computers via LimeSurvey software (Version 2.06; Schmitz, 2015). The stimuli used were 48 English-Spanish word pairs taken from Metcalfe (2002). The word-pairs selected were classified as being either easy (16 pairs), medium (16 pairs) or difficult (16 pairs) to learn (as indicated by Metcalfe, 2002). For the purpose of phases 2 and 4 (see Figure1), these were divided between two lists, thus each list contained eight word-pairs of each difficulty level.

Participants' implicit theories of intelligence were measured using Dweck's (1999) Theories of Intelligence scale. The scale contains 8 items, requiring participants to rate on a scale from 1(strongly agree) to 6(strongly disagree) the extent to which they endorse statements about the nature of intelligence. Four of these statements promote an entity view of intelligence (eg. *"To be honest, you can't really change how intelligent you are"*), and four endorse an incremental view (eg. *"No matter who you are, you can significantly change your intelligence level")*. The incremental statements were reverse coded, so that each person received a mean theories of intelligence score between 1 and 6, whereby scores closest to 1 indicate entity orientation, and scores closer to 6 reflect an incremental orientation. Although various forms of the scale have been used in the literature (eg Dweck, 1990; Claro et al., 2016), it is consistently reported as having good reliability (internal and 2-week test-retest) and validity (see Dweck et al., 1995; Blackwell et al., 2007; De Castella & Byrne, 2015). Miele and colleagues (2011) reported the full 8 item scale as having good internal reliability ($\alpha = .96$), which the findings from the present study were consistent with ($\alpha = .89$).

Participants also completed a demographics and language experience questionnaire designed to screen for proficiency in Spanish (see Appendix D).

Procedure

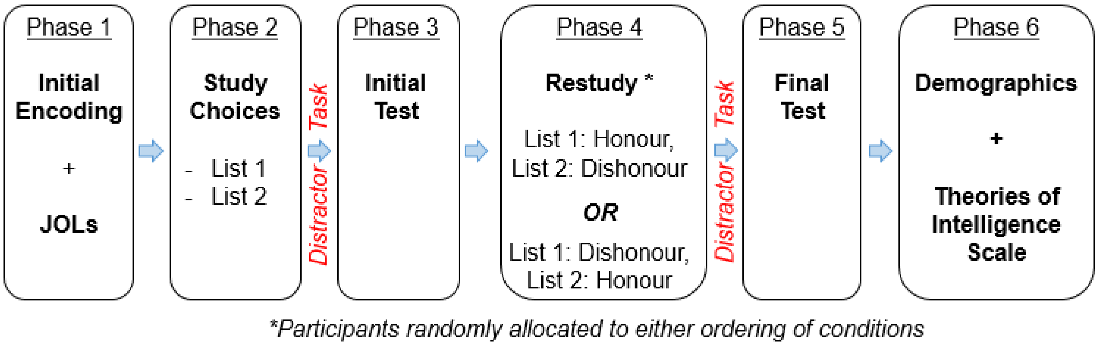

Figure 1 Experiment proper procedure.

Participants were informed verbally that they would be asked to learn some English words and their Spanish translations. They were told that they would be presented with the words for fixed amounts of time, asked to make judgements about their learning, restudy the material, and complete some tests along the way. Participants were asked to try to learn the words to the best of their ability, however they were informed they could skip test items that they were unable to answer. All other instructions, which were provided onscreen by the computer throughout the study, can be found in Appendix E.

The study began with demonstration slides to familiarise participants with the layout of the word-pairs (English word on the left, Spanish translation on the right), and to allow them to practise using the JOL rating scale.

The experiment proper comprised 6 main phases (see Figure 1), and was designed to follow the basic structure of a student preparing for an exam. Phase one was an initial encoding and JOL phase. As in Kornell and Metcalfe's studies (2006, 3a & 3b), the English-Spanish word-pairs were presented individually at a fixed rate, and participants were asked to make a JOL immediately after each word-pair disappeared. In Kornell and Metcalfe's study, word-pairs were presented for 4 seconds each. Miele and colleagues (2011, Study 2) also used this same pace, however they concluded that this timing limited their participants' ability to make attributions based on engagement. It was therefore decided in the present study to extend the presentation time for each word-pair to 10 seconds, to increase the opportunity for participants to engage with each word-pair, whilst still limiting their ability to achieve mastery.

After the 10 seconds per word-pair expired, participants were asked to make a JOL. On each JOL screen, participants were asked to respond to the question *"How confident are you (0%-100%) that you could type the Spanish translation if tested in a few minutes?"* by clicking and dragging slider handles on a scale. The JOL screens also presented the instruction to *"Please answer quickly"* to discourage participants from using this time to rehearse or study the word-pair they had just viewed. After participants worked through the first list (List 1), a 30 second break was enforced to reduce fatigue effects. A countdown timer displayed onscreen informed participants of the break time remaining. Once the break concluded, participants worked through the second list (List 2) of word-pairs in the same manner as the first.

In the second phase, participants were asked to make study choices. As in Kornell and Metcalfe (2006, Study 3b), the English cue-words from the word-pairs shown in List 1 were presented simultaneously onscreen, and participants were instructed to select half of the words for restudy. In order to explore the extent to which participants could discriminate between learned and unlearned items, participants were required to choose from all the word-pairs presented in each list, and were instructed to *'select word-pairs to study that you think you have not yet learned'* (whereas Kornell and Metcalfe only allowed participants to choose from word-pairs they had not correctly recalled on an initial test). Therefore, participants were instructed to choose 12 out of 24 items from each list. Once participants completed their study choices for List 1, they were required to do the same for List 2 on the next screen. Before moving onto the initial test phase, participants completed a three minute maths problems distractor task.

The third phase was the initial cued-recall test. The instructions indicated that this was a self-paced practice test, for which only words spelled accurately would be marked as being correct. All of the English cue-words were presented individually in a randomized order, and participants were required to type the Spanish translation of each word. As in Kornell and Metcalfe (2006, Studies 3a & 3b), answers had to be spelled exactly as they were presented to be considered accurate.

In the fourth phase participants were able to restudy word-pairs based on the selections they made in phase three. The restudy phase had two within-participant conditions: honour, and dishonour. In the honour condition, each word-pair selected for restudy was shown on screen for 10 seconds. In the dishonour condition, only word-pairs that participants did not select for restudy were presented. If participants have accurate metacognitive monitoring, and make optimal study choices, their recall should be higher for their honour list than for their dishonour list. Therefore, participants with poorer metacognitive monitoring may benefit to a lesser extent from being allowed to choose the items they requested. Each participant completed both conditions, and was randomly allocated to either receive the honour condition for List 1 and the dishonour condition for List 2, or to receive the dishonour condition for List 1 and the honour condition for List 2. The instructions informed participants that *'The word pairs that appear may, or may not, be the ones you requested'*. This instruction was included to reduce the likelihood of participants becoming confused by being shown words that they had not requested, or assuming that the software was malfunctioning. Participants then completed another three minute maths problems filler task before moving onto phase five.

The fifth phase was the final test. Aside from being introduced as a final test, it was otherwise identical to the initial test in phase three (although the word orders were re-randomized).

In the final phase, participants filled out the demographics and language experience questionnaire, as well as the Theories of Intelligence Scale (Dweck, 1999).

Results

To assess the relationship between confidence (JOLs) and accuracy on the initial test, three measures of calibration were utilised: the calibration statistic, the over/under-confidence statistic (O/U), and the Adjusted Normalized Discrimination Index (ANDI). Calibration approaches have been advocated in the confidence-accuracy literature, as they are argued to be more informative and less biased than other measures, such as point-biserial correlations (e.g. Juslin, Olsson, and Winman 1996; Palmer, Brewer, Weber & Nagesh, 2013). Whilst each of these measures make assessments of the confidence-accuracy relationship, they examine discrete elements of it (Yaniv, Yates & Smith, 1991; Palmer et al., 2013). *The calibration statistic* provides information about how far accuracy deviates from the expected pattern of results that would occur if each level of confidence perfectly predicted accuracy. Figure (2) provides a visual representation of what perfect calibration would look like: 30% of the responses that are given 30% confidence are actually correct, 60% of the responses at 60% confidence are correct and so forth. The calibration statistic indexes how far observed calibration differs from the perfect calibration line, producing values between 0 (prefect calibration) to 1. *The O/U statistic* is a measure reflecting how average confidence ratings compare to overall accuracy, with values ranging from -1 (indicating on average, confidence was lower than accuracy) to +1 (indicating on average, confidence was greater than accuracy). *The ANDI statistic* is a measure of discrimination, revealing how informative confidence judgements are for discriminating between correct and incorrect judgments, with values ranging from 0 to 1 (perfect discrimination).

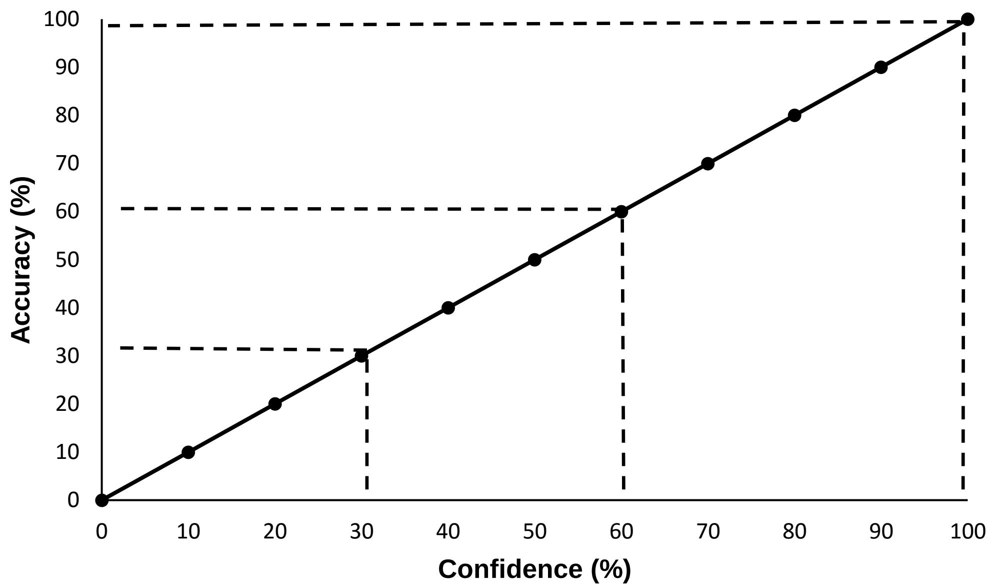

Figure 2. Perfect calibration curve. Each level of confidence accurately predicts accuracy.

To assess metacognitive monitoring during the study choice phase, an effect size (Cohen's *w*) was calculated for each participant from a 2 x 2 Chi-square contingency table, comparing study decisions and performance on the initial test. As a guide, Cohen's (1988) recommended values of 0.1, 0.3 and 0.5 were used as approximate guides to assess small, medium and large effects, respectively. Figure 3 depicts the contingency table for a participant with perfect metacognitive monitoring; all responses would fall in quadrant *a* (items not chosen for restudy and correctly recalled on initial test) and *d* (items chosen for restudy and not recalled on initial test).

		Accuracy on Initial Test	
		Correct	Incorrect
Restudy Decision	Do not restudy	*a* ✓	*b*
	Restudy	*c*	*d* ✓

Figure 3. Representation of 2 (accuracy) x 2 (restudy decision) Chi-square table used to calculate an effect size for each participant's metacognitive monitoring. All responses for perfect monitoring would fall in quadrants a and d.

Following Miele and colleagues (2011), mixed Analyses of Covariances (ANCOVAs) were used to assess the relationships between word-pair difficulty (with-in subjects, 3 levels: easy, moderate, difficult) and confidence-accuracy statistics, with theory of intelligence (between subjects) as the covariate. This approach allows an interaction term between difficulty and theory of intelligence to be obtained, which is of central interest to the current study's hypotheses[1]. Also following Miele et al., where significant interactions were found, they were followed up with bivariate correlations, to assess the relationship at each level of difficulty.

[1] As a covariate, TOI was not observed as having a main effect on any measure of metacognitive monitoring (all $F < 0.86$, all $p > .357$). As these effects were not related to the hypotheses, they are not discussed further.

As is common in the literature (eg. Robins & Pals, 2002; Miele et al., 2011), theory of intelligence was measured and entered as a continuous variable in all analyses, whereby smaller values indicate entity endorsement, and greater values suggest a more incremental orientation. This approach avoids the loss of power typically associated with dichotomising continuous variables (Cohen, 1983).

Before formal analyses were conducted, all data was screened to assure that appropriate assumptions had been met. Due to skew in the distribution of the calibration data, a log10 transformation was used. Inspection of the histograms, P-P plots and skewness statistic for each level of difficulty of the calibration statistic indicated the transformation was successful at reducing the bias, thus the transformed data was retained for analyses. In cases where the assumption of sphericity was violated, as indicated by a significant Mauchley's value, Greenhouse-geisser corrections were applied. All other assumptions were met.

Manipulation check

To confirm the anticipated difficulty difference between easy, moderate and difficult word-pairs, a one-way repeated measures Analysis of Variance (ANOVA) was conducted, with initial test performance as the dependent variable. Mauchley's test indicated a violation of sphericity, thus a Greenhouse-geisser correction was used. A main effect of difficulty was observed, $F(1.38, 88.19) = 228.07$, $p <.001$. Bonferroni adjusted paired-samples t-tests revealed participants recalled more easy words ($M = 6.29$, $SD =3.21$) than moderate word-pairs ($M = 1.12$, $SD =1.72$), $t(64) = 17.21$, $p <.001$, $95\%CI_{difference}[4.57, 5.77]$, $d = 2.48$, and more moderate words than difficult words ($M = 0.2$, $SD = 0.12$), $t(64) = 5.18$, $p < .001$, $95\%CI_{difference}[0.68, 1.53]$, $d = 0.64$., thus the manipulation was successful.

Confidence (JOLs) and accuracy

To assess the relationship between theory of intelligence and the utility of JOLs for predicting actual performance at each level of confidence, the effect of difficulty (within subjects) and theory of intelligence (covariate) on the calibration (transformed) statistic was analysed using an ANCOVA[2]. It was anticipated that greater endorsement of incremental views of intelligence would be associated with poorer calibration, especially for easy and difficult items (as perceived fluency was anticipated to be highest and lowest for these word-pairs, respectively). A significant main effect of difficulty was found, $F(1.58, 97.89) = 6.69$, $p = .002$. $\eta^2_p = .10$ (Greenhouse-Geisser correction). This main effect was interpreted in the context of the marginally significant interaction observed between difficulty and theory of intelligence, which was of central concern to the hypothesis, $F(1.58, 97.89) = 2.88$, $p = .073$, $\eta^2_p = .04$. To further investigate the nature of this interaction, a series of bivariate correlations were run to examine the relationship between theory of intelligence and calibration statistics for each level of word-pair difficulty. The relationships between theory of intelligence and calibration statistics were non-significant for both the easy ($r = -.085, p = .503$) and moderate ($r = .046, p = .714$) items. Whilst also non-significant at the alpha level of .05, the relationship observed for the difficult items trended in a positive direction, suggesting that for these items, endorsing a more incremental view of intelligence was associated with poorer calibration ($r = .235, p = .062$). Thus this finding offers some support for the hypothesised compromised metacognitive monitoring of incremental theorists, as observed by Miele et al., (2011).

[2] Descriptive statistics for confidence-accuracy analyses are reported in Table1

Table 1

Means (and Standard Deviations) for Confidence-Accuracy Statistics

	Easy	Moderate	Difficult	Total
Calibration Statistic*	0.15 (0.13)	0.15 (0.10)	0.08 (0.08)	0.11 (0.08)
Log10 Calibration Statistic*	-0.95 (0.36)	-0.98 (0.47)	-1.35 (0.54)	-1.06 (0.33)
O/U*	0.23 (0.23)	0.28 (0.17)	0.20 (0.14)	0.23 (0.14)
ANDI	-	-	-	0.32 (0.20)
Cohen's w	0.29 (0.20)	0.19 (0.21)	0.16 (0.14)	0.32 (0.21)

Note: Perfect metacognitive monitoring for statistics marked with * equals 0; otherwise, higher values represent superior monitoring. ANDI values were not calculated separately for easy, medium and difficult levels.

In order to investigate the effect of theory of intelligence on participants' average confidence ratings for each level of word-pair difficulty, an ANCOVA was conducted with the O/U statistic as the dependant variable. It was hypothesised that greater endorsement of incremental views of intelligence would be associated with under-confidence for easy word-pairs, and over-confidence for difficult word-pairs. The ANCOVA revealed a non-significant main effect of difficulty, $F(1.39, 87.40) = 1.03$, $p = .336$, $\eta^2_p = .02$ (Greenhouse-Geisser correction). The interaction between difficulty and theory of intelligence was also non-significant, $F(1.39, 87.40) = 0.95$, $p = .362$, $\eta^2_p = .02$. These findings suggest that on average, the extent of over- or under-confidence did not differ as a result of incremental indorsement and difficulty levels.

To assess the relationship between theory of intelligence and the utility of participants' JOLs in discriminating between learned and unlearned word-pairs, a linear regression was run between theory of intelligence and mean ANDI scores[3]. It was hypothesised that greater endorsement of incremental views of intelligence would be associated with poorer discrimination between learned and unlearned word-pairs. Theory of intelligence was not found to predict a significant amount of the variance in participants' ANDI statistic, $R^2 = .01$, $F(1, 63) = .67$, $p = .415$. Neither the unstandardized (B = -.02) or the standardized (β = -.10) slopes differed significantly from 0, $t(63) = -.82$, $p = .415$. Consequently, theory of intelligence did not appear to influence overall discrimination.

Study decisions & accuracy

In order to investigate whether theories of intelligence impact metacognitive monitoring after a delay (i.e. when making study decisions), an ANCOVA with each participant's metacognitive monitoring effect size (Cohen's *w*) as the dependant variable was conducted. Two possible outcomes were predicted: either incremental theorists would have difficulty discerning the items that they had and had not learned, and thus would have smaller effect sizes for their metacognitive monitoring, or alternatively, the delayed-JOL-effect (Nelson & Dunlosky, 1991) would result in any differences between incremental and entity theorists disappearing. The ANCOVA revealed a main effect of difficulty, $F(2, 126) = 9.72$, $p < .001$, $\eta^2_p = .13$. Of central interest to the present study, an interaction was observed between difficulty and theories of intelligence, $F(2,126) = 5.59$, $p = .005$, $\eta^2_p = .08$. To investigate the nature of the interaction, separate bivariate correlations were conducted for each level of difficulty. A significant weak-moderate negative

---

[3] Calculation of the ANDI statistic requires participants to provide at least one correct and one incorrect response. The majority of participants did not achieve this for moderate and difficult items, hence ANDI scores were collapsed across all difficulty levels for this analysis.

relationship was observed for easy word-pairs, r = -.260, p = .036, suggesting that greater endorsement of incremental views of intelligence was associated with poorer metacognitive monitoring during study decisions. The correlations for moderate word-pairs (r =.049) and difficult word-pairs (r = .202) were both non-significant, (p = .699 and, p = .106, respectively). These findings provide some evidence that incremental theorists may experience compromised metacognitive monitoring, even with delayed judgments.

Study approaches

What did people choose? To investigate the study approach participants tended to follow (DR vs. RPL), a paired-samples t-test was conducted between mean JOLs for items chosen for restudy and the items not chosen. As participants were expected to follow a RPL approach (by choosing word-pairs to restudy that they did not currently know, but perceived would be easy to learn), it was hypothesised that mean JOLs would be greater for the items chosen for restudy than those not chosen. Contrary to this hypothesis, participants chose items for restudy that had lower JOLs (*M = 37.47, SD = 13.82*) than items not chosen for restudy (*M = 48.66, SD = 16.94*), which represented a medium effect ($t(64) = -5.37$, $p <.001$, 95% CI_{diff}[-15.36, -7.03], d = .72). Furthermore, an ANCOVA investigating the effect of choice on JOLs, with theory of intelligence as the covariate, revealed a non-significant main effect for theory of intelligence, $F(1, 63) = 0.20$, $p = .656$, $\eta^2_p = .00$[4]. Thus, theory of intelligence did not appear to influence the study approach adopted. To allow a closer comparison to Kornell and Metcalfe's (2006, Study 3a & 3b) design, the same

[4] The ANCOVA produced a non-significant main effect of choice, $F(1, 63) = 0.56$, $p = .457$, $\eta^2_p = .00$. Given the moderate effect size revealed by the paired-samples t-test, the t-test analysis was considered as better representing the data.

analysis was run, however only for the word-pairs which participants got incorrect on the initial test. Once again, participants chose word-pairs for restudy that they had assigned lower JOLs to (M = 35.67, SD = 13.90) compared to the items not chosen for restudy (M = 41.32, SD = 15.98), which represented a small effect, $t(64) = -3.12$, $p = .003$, 95% CI_{diff}[-9.28, -2.04], $d = .38$. Moreover, non-significant main effect of theory of intelligence was again observed from an ANCOVA, $F(1, 63) = 1.64$, $p = .205$, $\eta^2_p = .03$[5]. Together, these results suggest participants in the present study tended to follow a DR approach.

What should people have chosen? To investigate which study approach (DR or RPL) was the more optimal approach in the current study, a linear regression was conducted. Based on the RPL framework, it was hypothesised that there would be a positive relationship between the JOLs for items chosen for restudy and final recall performance[6,7]. The regression revealed participants' mean JOLs for items chosen for restudy did not predict a significant amount of the variance in final test recall, $R^2 = .02$, $F(1, 62) = 1.33$, $p = .253$. Neither the unstandardized (B = 0.03) or the standardized (β = 0.15) slopes differed from 0, $t(62) = 1.15$, $p = .253$. Once more, to allow a closer comparison to Kornell and Metcalfe (2006, Study 3a & 3b), the relationship of JOLs chosen only for the items that were incorrect on the initial test were used to predict final test recall (for these items). Again, participants' mean JOLs for items that they chose for restudy (and got wrong on an initial test) did not predict a significant amount of the variance in their final test recall, $R^2 = .02$, $F(1, 62)$

[5] Once again, the ANCOVA produced a non-significant main effect of choice, $F(1, 63) = 0.46$, $p = .457$, $\eta^2_p = .01$, despite the moderate effect size found in the t-test, thus the t-test is considered to best represent the data..
[6] These analyses were only conducted on the final recall of items in the honour condition.
[7] Due to a software error, one participant's final recall data was not saved; analyses for final recall are therefore based off 64 participants.

= 1.52, p = .223. Neither the unstandardized (B = 0.02) or the standardized (β = 0.16) slopes differed from 0, t(62) = 1.23, p = .223. These results suggest that the current participants' performance on the final test was not associated with their tendency to select items based off a RPL or DR framework.

Finally, regardless of the study approach taken, it was also of interest to know if participants were generally able to make optimal study decisions. To investigate this, a paired-samples t-test was conducted between final test performance in the honour and dishonour conditions. If participants made good study choices, their recall for the word-pairs in the list in which their study decisions were honoured would be expected to be greater than their recall for the dishonour condition. Surprisingly, this was not the case. Participants' recall for word-pairs in the honour condition (M = 5.25, SD = 3.51) was no different than their recall in the dishonour condition (M = 5.25, SD = 3.37), t = 0, thus it appears they did not benefit at all from making study decisions.

Discussion

The current study sought to investigate the relationship between theories of intelligence and metacognitive monitoring. Contrary to the hypotheses, most of the confidence-accuracy measures failed to reveal impaired monitoring for incremental theorists. Greater endorsement of an incremental theory of intelligence was not associated with general over-confidence or under-confidence, nor did it seem to hinder participants' ability to immediately discriminate learned from unlearned word-pairs, for any level of difficulty. The only support for the hypotheses regarding immediate confidence and accuracy was a trend towards a positive relationship between theories of intelligence and the calibration statistic for difficult word-pairs. Thus, the current findings are largely in contrast to those of Miele and colleagues

(2011, Study 1; & Molden, 2010, Study 4). Impaired metacognitive monitoring was observed however for incremental theorists in their ability to appropriately choose easy items for restudy. Although it was anticipated that greater incremental endorsement may be associated with poorer discrimination in study choices, this finding was somewhat surprising given the very limited indication of monitoring impairment for incremental theorists in their initial confidence ratings.

On most measures of metacognitive accuracy, the current study failed to replicate the impairments for incremental theorists observed by Miele et al. (2011). At best, only limited support was found for the suggestion that theories of intelligence influence metacognitive monitoring. These findings may therefore imply that less concern about the possible drawbacks of incremental orientation is warranted than Miele and colleagues (Miele & Molden, 2010; Miele et al., 2011; Miele, Son & Metcalfe, 2013) have suggested.

When considering the contrast in results observed between Miele et al.'s (2011) findings and the present study, it is important to consider other factors that may have contributed to the disparity. One such example may be linked to the distribution of scores in the present study on theories of intelligence scale. There are two possible explanations of how the spread of data could reduce the likelihood of observing strong relationships between theory of intelligence and metacognitive monitoring. Firstly, if the distribution of scores was highly clustered, this would undermine the ability of the correlation statistic to detect strong relationships. As this was not apparent in the current sample, with mean scores ranging from 2 to 6, it is unlikely that the correlations were statistically undermined. A more likely explanation is that not enough participants' mean theories of intelligence score fell on either side of the scale's midpoint. Indeed, only 6 out of 65 participants had a mean score less than the midpoint of 3 (indicating they endorsed more entity views of intelligence), whilst 59 participants' score fell above 3 (suggesting more incremental orientation). Thus, the current study was mostly commenting on participants who ultimately endorsed incremental views of intelligence, just to a greater or lesser extent. If the midpoint of the scale represents a critical threshold, whereby having a mean theories of intelligence score above 3 makes one's interpretive frameworks fundamentally different to those whose mean score falls below 3, then the failure to observe impaired monitoring for incremental theorists may be due to an inadequately sized group of entity theorists in the current study.

The recruitment of a sample with such a small proportion of entity theorists was unanticipated, as similar results have either not been reported in studies of theories of intelligence (e.g. Miele et al., 2011; Miele & Molden, 2010), or else the magnitude of the difference between incremental theorists and entity theorists recruited was not as extreme (e.g. Hong et al., 1999). For example, even though only 30 out of 97 of participants recruited by Hong and colleagues in Study 1, and 64 out of 168 participants from Study 2, were classified as entity theorists, the percentages of entity theorists recruited from their university population (31% and 38% respectively) were far greater than the 9% in the current study. The relatively small proportion of entity theorists observed may have occurred due to chance, thus with additional recruitment more entity theorists may participate, which might allow for a clearer comparison of the two theories of intelligence.

Alternatively, the design of the current study (in which participants were asked to learn the word-pairs to the best of their ability over multiple study trials, and with multiple tests, prior to completing the theory of intelligence scale) may have primed them to provide more incremental responses. Given that theories of intelligence were assessed in the same way by Miele and colleagues (2011), it seems unlikely that completing the scale after studying would bias mean ratings. To address the possibility that a discreet element of the current study (not present in Miele et al.,) influenced ratings, future research could utilise a pre-testing session to assess theories of intelligence (as in Miele & Molden, 2010, Studies 3-5). It should also be noted that some researchers have suggested it may be preferable at times to assess theories of intelligence with entity endorsing items alone (e.g. Dweck, 1999; Hong et al., 1999). Hong and colleagues reported finding evidence in pilot studies that participants who endorse entity statements also tend to demonstrate increased incremental endorsement as they work through the scale. This observation may be due to the perceived social desirability of incremental orientation (Hong et al., 1999). Despite this suggestion however, using the full 8-item theories of intelligence scale, Miele and colleagues (& Molden, 2010; et al., 2011) still found differences between entity and incremental theorists.

Another possible explanation for why strong relationships were not observed between theories of intelligence and metacognitive monitoring may be as a result of the rate of stimulus presentation enforced in the current study. Indeed, Miele and colleagues' (et al., 2011, Study 1; & Molden, 2010; Studies 4 &5) findings of compromised monitoring for incremental theorists occurred when participants were able to pace their own study. Moreover, in Miele et al.'s (2011) Study 2, in which items were presented for 4 seconds each, incremental theorists' confidence ratings did not differ between fluency conditions. Miele and colleagues suggested this may have been the result of the study pace limiting incremental theorists' ability to feel engaged in the material. Whilst this may provide a plausible explanation in that particular study, in the current study this explanation seems less likely to apply, due to the 10 second rate of presentation. Considering incremental theorists in Miele and colleagues' Study 1 spent a mean time of 3.46 seconds studying each Indonesian-English word-pair (whilst entity theorists spent 5.59 seconds), and demonstrated poor calibration across word-pair difficulty, it seems unlikely that 10 seconds per word-pair in the present study would have comparatively impaired their ability to feel engaged with the material.

Another aim of the present study was to investigate the study approaches that participants tend to follow. Specifically, the present study aimed to add further support to the RPL paradigm (Kornell & Metcalfe, 2006; Metcalfe & Kornell, 2005). Given the time constraints imposed by the fixed stimulus presentation pace in current study, it was surprising to observe that participants' study choices reflected a DR approach. This finding was in direct contrast to the anticipated pattern of results, as the RPL literature suggests when study time is limited, people focus their study resources on the easiest unlearned items (Kornell & Metcalfe, 2006; Metcalfe & Kornell, 2005; Son & Metcalfe, 2000). Although the pace of 10 seconds per word-pair provided greater study time than participants in Kornell and Metcalfe's Studies 3a and 3b were allowed (4 seconds per word-pair), mastery was still not likely to be obtained. Indeed, mean recall on both the initial and final tests were quite low (15.63% and 22.89% respectively), as were mean JOLs, (M = 42.62, SD = 28.91), thus not only was mastery not likely, participants also appeared to be aware of this. Moreover, this pattern of study allocation was observed when analyses were restricted to choices of items that were demonstrated as being unknown by performance on the initial test. A likely explanation for this surprising finding is that the instructions provided during the study choice phase were interpreted by participants as requesting choices be made in a DR manner. The instruction to '*select* *word-pairs to study that you think you have not yet learned'* was designed to assess the central concern of the present study: whether incremental theorists would exhibit compromised metacognitive monitoring, in this case, when discerning known from unknown items after a delay. It is possible however that this instruction may have encouraged participants to select the items they were *most* certain they had not learned (which would likely be the word-pairs assigned the lowest JOLs – thus resulting in a DR pattern of choosing). Therefore, any conclusions made in regards to typical study behaviours based on the findings of the present study should be tentative, and must take this possible procedural confound into account.

Another possible explanation for the observed DR approach in the present study may again result to the predominantly incremental sample. As incremental theorists are oriented towards challenge and learning goals (Dweck, 1999), they may have intentionally sought the most difficult items. Again, the nature of the current sample limits the opportunity to meaningfully comment on this possibility, therefore future research may benefit from recruiting more entity theorists, and utilising alternate instructions or designs to investigate whether theory of intelligence influences preferred study behaviours.

Surprisingly, in the current study, participants did not seem to benefit from controlling their own study; neither a DR or RPL study approach, nor controlling study in general, was associated with higher final test recall. It is possible that this finding may also have occurred as a consequence of the current study's design. As participants tended to choose word-pairs with low JOLs, in the honour condition they would have received these to study, however in the dishonour condition they would have tended to receive word-pairs with higher JOLs (more akin to a RPL approach). Based on Kornell and Metcalfe's (2006, Studies 3a & 3b) research indicating that following a RPL results in higher final test recall than a DR approach, the benefits they also observed on recall from participants being able to control their own study (compared to having their study choices dishonoured) may have been cancelled out in the current study by the high-JOL-nature of the dishonour condition.

In summary, whilst the current study found some indications of compromised metacognitive monitoring for incremental theorists, the majority of measures failed to demonstrate this effect. The present study therefore does not provide compelling support for the suggestion that promoting incremental orientation is likely to impair student's metacognitive accuracy (Miele & Molden, 2010; Miele et al., 2011; Miele, Son & Metcalfe, 2013). However, in light of the aforementioned limitations (especially the predominantly incremental sample), any conclusions based on the current study can only be tentative at best. If future research addresses these limitations and finds further evidence for the impairment of incremental theorists' metacognitive monitoring, this may warrant reconsideration of the advice educators provide students (Miele, et al., 2013). Currently, researchers are calling for policy makers and educators to nurture incremental orientations in students through educational systems and interventions (Rattan, Savani, Chugh & Dweck, 2015). To be clear, given the literature identifying the overwhelming benefits of being an incremental theorist on educational outcomes (e.g. Claro et al., 2016; Henderson & Dweck, 1990), it is not suggested that these interventions be opposed. Rather, if more compelling evidence indicates that incremental theorists experience impaired metacognitive monitoring, these interventions could perhaps be optimised by including strategies to help students compensate for any compromised monitoring they may experience. In such a case, incremental theorists may benefit from utilising strategies such as self-testing, which has been demonstrated not only as a strategy useful for assessing actual learning, but also provides an additional learning experience in itself (Finn & Tauber, 2015). Overall, whilst the aforementioned limitations mean the findings of the current study should be regarded as tentative, the findings provide little reason to suggest incremental theorists are likely to experience major metacognitive drawbacks during learning and study.

References

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, *38*(2), 113-125. doi: 10.1006/jesp.2001.1491

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child development*, *78*(1), 246-263. doi: 10.1111/j.1467-8624.2007.00995.x

Claro, S., Paunesku, D., & Dweck, C. S. (2016). Growth mindset tempers the effects of poverty on academic achievement. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(31), 8664-8668. doi:10.1073/pnas.1608207113

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, *7*, 249-253. Retrieved from http://apm.sagepub.com/

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, N.J.: L. Erlbaum Associates, 1988.

De Castella, K., & Byrne, D. (2015). My intelligence may be more malleable than yours: The revised implicit theories of intelligence (self-theory) scale is a better predictor of achievement, motivation, and student disengagement. *European Journal of Psychology of Education*, *30*(3), 245-267. doi: 10.1007/s10212-015-0244-y

Diener, C. I., & Dweck, C. S. (1980). An analysis of learned helplessness: II. The processing of success. *Journal of Personality and Social Psychology, 39*(5), 940-952. doi:10.1037//0022-3514.39.5.940

Dunlosky, J., & Thiede, K. W. (1998). What makes people study more? An evaluation of factors that affect self-paced study. *Acta Psychologica*, *98*(1), 37-56. Sourced from: http://www.journals.elsevier.com/acta-psychologica/

Dweck, C.S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.

Dweck, C. S., Chiu, C. Y., & Hong, Y. Y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry*, *6*(4), 267. doi: 10.1207/s15327965pli0604_1

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*(2), 256-273. doi: 10.1037/0033-295X.95.2.256

Finn, B., & Tauber, S. K. (2015). When confidence is not a signal of knowing: How students' experiences and beliefs about processing fluency can lead to miscalibrated confidence. *Educational Psychology Review*, *27*(4), 567-586. doi: 10.1007/s10648-015-9313-7

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906-911. doi: 10.1037/0003-066X.34.10.906

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat.

Journal of Applied Developmental Psychology, *24*(6), 645-662. doi: doi:10.1016/j.appdev.2003.09.002

Henderson, V., & Dweck, C. S. (1990). Achievement and motivation in adolescence: A new model and data. In S. Feldman, & G. Elliott (Eds.), *At the threshold: The developing adolescent* (pp. 308–329). Cambridge, MA: Harvard University Press.

Hong, Y. Y., Dweck, C. S., Chiu, C. Y., Lin, D. M. S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, *77*(3), 588-599. doi:10.1037/0022-3514.77.3.588

Juslin, P., Olsson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can be inferred from the low confidence–accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1304-1316. doi: 10.1037/0278-7393.22.5.1304

Koriat, A. (2008). Easy comes, easy goes? The link between learning and remembering and its exploitation in metacognition. *Memory & Cognition*,*36*(2), 416-428. doi: 10.3758/MC.36.2.416

Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of memory and Language*,*52*(4), 478-492. doi:10.1016/j.jml.2005.01.001

Kornell, N., & Bjork, R. A. (2008). Optimising self-regulated study: The benefits—and costs—of dropping flashcards. *Memory*, *16*(2), 125-136. doi: 10.1080/09658210701763899

Kornell, N., & Metcalfe, J. (2006). Study efficacy and the region of proximal learning framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(3), 609-622. doi: 10.1037/0278-7393.32.3.609

Masur, E. F., McIntyre, C. W., & Flavell, J. H. (1973). Developmental changes in apportionment of study time among items in a multitrial freerecall task. *Journal of Experimental Child Psychology, 15,* 237–246. doi: 10.1016/0022-0965(73)90145-8

Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning? *Journal of Experimental Psychology: General*, *131*(3), 349-363. doi: 10.1037//0096-3445.131.3.349

Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language*, *52*(4), 463-477. doi: doi:10.1016/j.jml.2004.12.001

Miele, D. B., Finn, B., & Molden, D. C. (2011). Does easily learned mean easily remembered? It depends on your beliefs about intelligence. *Psychological Science*, *22*(3), 320-324. doi: 10.1177/0956797610397954

Miele, D. B., & Molden, D. C. (2010). Naive theories of intelligence and the role of processing fluency in perceived comprehension. *Journal of Experimental Psychology: General*, *139(*3), 535-557. doi: 0.1037/a0019745

Miele, D. B., Son, L. K., & Metcalfe, J. (2013). Children's naive theories of intelligence influence their metacognitive judgments. *Child Development*,*84*(6), 1879-1886. doi: 10.1111/cdev.12101

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The “delayed-JOL effect”. *Psychological Science, 2*(4), 267-270. doi: 10.1111/j.1467-9280.1991.tb00147.x

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the" labor-in-vain effect.". *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(4), 676-686. doi: 10.1037/0278-7393.14.4.676

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and some new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 125–173). San Diego, CA: Academic.

Palmer, M. A., Brewer, N., Weber, N., & Nagesh, A. (2013). The confidence-accuracy relationship for eyewitness identification decisions: Effects of exposure duration, retention interval, and divided attention. *Journal of Experimental Psychology: Applied*, *19*(1), 55-71. doi: 10.1037/a0031602

Robins, R., & Pals, J. (2002). Implicit self-theories in the academic domain: Implications for goal orientation, attributions, affect, and self-esteem change. *Self and Identity, 1*(4), 313-336. doi: 10.1080/15298860290106805

Schmitz, C. (2015). LimeSurvey: An open source survey tool (Version 2.06). Hamburg, Germany: LimeSurvey Project.

Schwartz, B. L., & Perfect, T. J. (2002). Introduction: Toward an applied metacognition. In T. J. Perfect & B. L. Schwartz, (Eds.), *Applied metacognition* (pp. 1-11). Cambridge, United Kingdom: Cambridge University Press.

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 204-221. doi: 10.1037//0278-7393.26.1.204

Yaniv, I., Yates, J. F., & Smith, J. K. (1991). Measures of discrimination skill in probabilistic judgment. *Psychological Bulletin*, *110*(3), 611-617. doi: 10.1037/0033-2909.110.3.611

Appendices

Appendix A: Ethics approval letter

Dear Dr Palmer

Ethics Ref: H0012660
Title: Confidence in memory

This email is to confirm that the following amendment was approved by the Chair of the Tasmania Social Sciences Human Research Ethics Committee on 17/5/2016:

- Addition of researchers Prof Andrew Heathcote, Dr Nicole McCallum, Ms Frances Parkes, and Dr Matthew Gretton.
- Addition of student researchers Valera Griffin, Laura Brumby, Terry Purton, and Caitlin Gleeson.
- Addition of Theories of Intelligence Questionnaire.
- Extension of data collection for a further three years.
- Revised Information Sheet for Participants.

All committees operating under the Human Research Ethics Committee (Tasmania) Network are registered and required to comply with the National Statement on Ethical Conduct in Human Research (NHMRC 2007, updated May 2015).

This email constitutes official approval. If your circumstances require a formal letter of amendment approval, please let us know.

Should you have any queries please do not hesitate to contact me.

Kind regards

Katherine

Katherine Shaw

Executive Officer, Social Sciences HREC
Office of Research Services | Research Division
University of Tasmania
Private Bag 1
Hobart TAS 7001
T +61 3 6226 2763
[www.utas.edu.au/research]www.utas.edu.au/research

UNIVERSITY of TASMANIA

CRICOS 00586B

Appendix B: Information Sheet

Locked Bag 1342 Launceston
Tasmania 7250 Australia
Phone (03) 6324 3004 Fax (03) 6324 3168
matthew.palmer@utas.edu.au

UTAS

Metacognition and Study Decisions Research

Information Sheet for Participants

UNIVERSITY OF TASMANIA

1. Invitation

My name is Laura Brumby, and I am a student at the University of Tasmania. I can show you my student ID card to verify this. I would like to invite you to participate in a psychology experiment about metacognition and study decisions. The experiment is being conducted under the supervision of Dr. Matthew Palmer of the Department of Psychology at the University of Tasmania.

2. *What is the purpose of this study?*

The experiment is investigating factors that affect people's memory for English-foreign language word-pairs, and their confidence judgements regarding their learning for these word-pairs.

3. Why have I been invited to participate?

For this experiment, we are looking for people aged 18 years or more who are able to fluently read English, and have normal or corrected to normal vision (i.e., glasses or contact lenses are fine). Participants must not be fluent in Spanish.

Participation in this study is voluntary – you are entirely free to choose to participate or not, and there will be no consequences if you decide not to participate. If you do participate, any information you provide will be anonymous and no participants in the experiment will be individually identifiable.

4. What will I be asked to do?

Participation would require approximately 50 minutes of your time on only one occasion and would take place in a room at the UTAS campus. The experiment involves learning English-foreign language word pairs, completing some recall tasks and confidence judgements on these word pairs, as well as filling out a brief questionnaire.

5. Are there any possible benefits from participation in this study?

The results of this study will help us to understand what factors affect people's memory and judgements of learning for English-foreign language word-pairs. This information will be useful to help inform research on optimal study behaviours. You would be reimbursed for your time with 1 research credit, or $20.

6. Are there any possible risks from participation in this study?

There are no foreseeable risks associated with participating in this study.

7. What if I change my mind during or after the study?

That's fine - you are free to withdraw from the study at any time, and without providing an explanation. If you choose to withdraw during the study, your responses will be destroyed. If you

Locked Bag 1342 Launceston
Tasmania 7250 Australia
Phone (03) 6324 3004 Fax (03) 6324 3168
matthew.palmer@utas.edu.au

UTAS

UNIVERSITY OF TASMANIA

complete the study, you will not be able to withdraw your data because it will be stored in anonymous form (and so we will not be able to identify which responses are yours).

8. What will happen to the information when this study is over?

The data from this study will be kept in secure storage on the University of Tasmania premises for a period of five years after any publications (e.g., in academic journals) that involve the data. After this period, the data will be archived. Only the researcher will have access to the raw data.

The data will be stored anonymously. All responses will be anonymous and no identifying information will be collected from participants.

9. How will the results of the study be published?

The results from this study will be published in an honours thesis. The results may also be published in an academic journal. Once the study has been completed, you will be able to access the results by visiting the website below:

http://www.utas.edu.au/psychology/research/research-project-reports

No individual participants will be identifiable in the publication of the results.

10. What if I have questions about this study?

If you have any questions about this study, please feel free to contact us via phone on (03) 6324 3004 (Matt Palmer) or by email: matthew.palmer@utas.edu.au

This study has been approved by the Tasmanian Social Sciences Human Research Ethics Committee. If you have concerns or complaints about the conduct of this study, please contact the Executive Officer of the HREC (Tasmania) Network on (03) 6226 7479 or email human.ethics@utas.edu.au. The Executive Officer is the person nominated to receive complaints from research participants. Please quote ethics reference number H0012660.

This information sheet is for you to keep. If you would like to participate in this study, please ask the researcher for a Consent Form to complete.

Thank you for your attention - your time is very much appreciated!

Appendix C: Consent Form

Locked Bag 1342 Launceston
Tasmania 7250 Australia
Phone (03) 6324 3004 Fax (03) 6324 3168
matthew.palmer@utas.edu.au

UTAS

U N I V E R S I T Y O F T A S M A N I A

Metacognition and Study Decisions Research

Participant Consent Form

1. I agree to take part in the research study named above.
2. I have read and understood the Information Sheet for this study.
3. The nature and possible effects of the study have been explained to me.
4. I understand that the study involves reading passages of text and answering questions about them.
5. I understand that participation involves no foreseeable risks.
6. I understand that all research data will be securely stored on the University of Tasmania premises for five years from the publication of the study results, and will then be destroyed unless I give permission for my data to be archived.

   I agree to have my study data archived. (Note that your data will be stored anonymously.)

   Yes ☐ No ☐
7. Any questions that I have asked have been answered to my satisfaction.
8. I understand that the researchers will maintain confidentiality and that any information I supply to the researcher will be used only for the purposes of the research.
9. I understand that the results of the study will be published so that I cannot be identified as a participant.
10. I understand that my participation is voluntary and that I may withdraw at any time without any effect.

    I understand that I will not be able to withdraw my data after completing the experiment as my data will be anonymous.

Participant's name: ____

Participant's signature: ____

Date: ____________________

Locked Bag 1342 Launceston
Tasmania 7250 Australia
Phone (03) 6324 3004 Fax (03) 6324 3168
matthew.palmer@utas.edu.au

UTAS

Statement by Investigator

☐ I have explained the project and the implications of participation in it to this volunteer and I believe that the consent is informed and that he/she understands the implications of participation.

If the Investigator has not had an opportunity to talk to participants prior to them participating, the following must be ticked.

☐ The participant has received the Information Sheet where my details have been provided so participants have had the opportunity to contact me prior to consenting to participate in this project.

Investigator's name: ______________________________

Investigator's signature: ______________________________

Date: ____________

U N I V E R S I T Y O F T A S M A N I A

Appendix D: Demographics and Language Experience Questionnaire

Participant Demographics

Age: _________ years

Gender: _________

1. Do you speak any languages, other than English?
   - ☐ No
   - ☐ Yes – Please specify which language(s) ____________________________
2. Have you ever studied any languages, other than English?
   - ☐ No
   - ☐ Yes – Please specify which language(s), and how long you studied them for

     ____
3. If you answered "Yes" to Question 2, please indicate how you studied each language (if you have studied more than one other language, please write next to the appropriate method which language you studied via this method

   – eg ☑ **At school** - *Spanish*)
   - ☐ At school
   - ☐ At University
   - ☐ Private lessons
   - ☐ Self-taught (using books, Apps etc)
   - ☐ Other: __________________

Appendix E: Study Instructions (via Limesurvey)

In this study you will be shown word-pairs containing English words and their Spanish translations.

Example word-pair: airport – aeropuerto

(English word) (Spanish word)

Your task is to learn these words to the best of your ability in preparation for a final test. In the final test, you will be asked to type the Spanish translation of each English word.

After each word pair is shown, you will be asked to rate how confident you are that in a few minutes you will be able to come up with the Spanish word when the English word is presented. You will be asked to make this rating on a scale between 0% - 100%, whereby 0% means you are not at all confident, and 100% means you are very confident that you will be able to type the Spanish translation correctly.

Example word-pair: airport – aeropuerto

Example confidence rating: How confident are you (0%-100%) that you could type the Spanish translation if tested in a few minutes? Please answer quickly. Please click and drag the slider handles to enter your answer.

When the study phase begins, each word pair will be shown for 10 seconds. When you are ready to begin the study phase, click next.

[List1]

It is now time for a quick break. In 30 seconds you will be shown more word pairs to learn.

[List2]

Below are some of the English cues from the word-pairs you were shown in Phase 1. In a moment, you will have the opportunity to study half of these words with their Spanish translations. In order to achieve the best possible result on the upcoming final test (in which you will have to type the Spanish translation of each English word), you should select word-pairs to study that you think you have not yet learned. Please choose EXACTLY TWELVE (12) of the words from the list below to restudy. *[Same instructions administered for List 2]*

[Filler task 1] Now you have 3 minutes to solve as many of the following maths problems as possible. Please enter your answers in the correct order in the box below. Separate each answer with a comma (,).

You will now be given a practice test on the word-pairs you were shown in Phase 1. When each English word is shown, your task is to type the Spanish translation of that word. Please take care with your spelling, as only words that are spelled correctly will be marked as being right. Once you have typed an answer, please click next to move on to the next item. Click next to begin the test.

[Cued-recall List 1]

[Cued recall List 2]

You will now be given the opportunity to study half of the word-pairs. The word-pairs will appear on the screen one at a time for 7 seconds each. The word pairs that appear may, or may not, be the ones you requested.

[Restudy List 1]

[Restudy List 2]

[Distractor task 2, instructions as above]

It is now time for the final test. When each English word is shown, your task is to type the Spanish translation of that word. Please take care with your spelling, as only words that are spelled correctly will be marked as being right. Once you have typed an answer, please click next to move on to the next item.

[Final Test]

[Demographics questionnaire]

[Theories of intelligence scale]

Thank you for your participation. Your time is very much appreciated. Please see the researcher if you have any comments or questions.

Appendix F: Output Summaries of Analyses

Manipulation Check

Descriptive Statistics

	Mean	Std. Deviation	N
E_Initial_Recall	6.2923	3.21482	65
M_Initial_Recall	1.1231	1.71854	65
D_Initial_Recall	.0154	.12403	65

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
difficulty	Sphericity Assumed	1459.200	2	729.600	228.074	.000	.781
	Greenhouse-Geisser	1459.200	1.378	1058.937	228.074	.000	.781
	Huynh-Feldt	1459.200	1.398	1043.497	228.074	.000	.781
	Lower-bound	1459.200	1.000	1459.200	228.074	.000	.781
Error(difficulty)	Sphericity Assumed	409.467	128	3.199			
	Greenhouse-Geisser	409.467	88.191	4.643			
	Huynh-Feldt	409.467	89.496	4.575			
	Lower-bound	409.467	64.000	6.398			

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	E_Initial_Recall	6.2923	65	3.21482	.39875
	M_Initial_Recall	1.1231	65	1.71854	.21316
Pair 2	M_Initial_Recall	1.1231	65	1.71854	.21316
	D_Initial_Recall	.0154	65	.12403	.01538

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	E_Initial_Recall - M_Initial_Recall	5.16923	2.42106	.30030	4.56932	5.76914	17.214	64	.000
Pair 2	M_Initial_Recall - D_Initial_Recall	1.10769	1.72412	.21385	.68048	1.53491	5.180	64	.000

Confidence (JOLs) and accuracy

Descriptive Statistics

	Mean	Std. Deviation	N
log10_CStat_Easy	-.9455	.35777	64
log10_CStat_Mod	-.9827	.47202	64
log10_CStat_Hard	-1.3499	.54147	64

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
difficulty	Sphericity Assumed	1.871	2	.935	6.692	.002	.097
	Greenhouse-Geisser	1.871	1.579	1.185	6.692	.004	.097
	Huynh-Feldt	1.871	1.639	1.141	6.692	.003	.097
	Lower-bound	1.871	1.000	1.871	6.692	.012	.097
difficulty * TOI	Sphericity Assumed	.805	2	.403	2.881	.060	.044
	Greenhouse-Geisser	.805	1.579	.510	2.881	.073	.044
	Huynh-Feldt	.805	1.639	.491	2.881	.071	.044
	Lower-bound	.805	1.000	.805	2.881	.095	.044
Error(difficulty)	Sphericity Assumed	17.334	124	.140			
	Greenhouse-Geisser	17.334	97.888	.177			
	Huynh-Feldt	17.334	101.633	.171			
	Lower-bound	17.334	62.000	.280			

Correlations

		TOI	log10_CStat_Easy
TOI	Pearson Correlation	1	-.085
	Sig. (2-tailed)		.503
	N	65	65
log10_CStat_Easy	Pearson Correlation	-.085	1
	Sig. (2-tailed)	.503	
	N	65	65

Correlations

		TOI	log10_CStat_Mod
TOI	Pearson Correlation	1	.046
	Sig. (2-tailed)		.714
	N	65	65
log10_CStat_Mod	Pearson Correlation	.046	1
	Sig. (2-tailed)	.714	
	N	65	65

Correlations

		TOI	log10_CStat_Hard
TOI	Pearson Correlation	1	.235
	Sig. (2-tailed)		.062
	N	65	64
log10_CStat_Hard	Pearson Correlation	.235	1
	Sig. (2-tailed)	.062	
	N	64	64

OU Stat

Descriptive Statistics

	Mean	Std. Deviation	N
OU_Easy	.23278846	.225598788	65
OU_Mod	.27625000	.174052962	65
OU_Hard	.19528846	.136765754	65

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
difficulty	Sphericity Assumed	.040	2	.020	1.034	.358	.016
	Greenhouse-Geisser	.040	1.387	.029	1.034	.336	.016
	Huynh-Feldt	.040	1.431	.028	1.034	.338	.016
	Lower-bound	.040	1.000	.040	1.034	.313	.016
difficulty * TOI	Sphericity Assumed	.037	2	.018	.949	.390	.015
	Greenhouse-Geisser	.037	1.387	.027	.949	.362	.015
	Huynh-Feldt	.037	1.431	.026	.949	.364	.015
	Lower-bound	.037	1.000	.037	.949	.334	.015
Error(difficulty)	Sphericity Assumed	2.444	126	.019			
	Greenhouse-Geisser	2.444	87.402	.028			
	Huynh-Feldt	2.444	90.162	.027			
	Lower-bound	2.444	63.000	.039			

ANDI

Descriptive Statistics

	Mean	Std. Deviation	N
ANDI_ALL	.31760180	.202267608	65
TOI	4.29038462	.869151125	65

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.103[a]	.011	-.005	.202785010

a. Predictors: (Constant), TOI

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.028	1	.028	.674	.415[b]
	Residual	2.591	63	.041		
	Total	2.618	64			

a. Dependent Variable: ANDI_ALL

b. Predictors: (Constant), TOI

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	.420	.128		3.293	.002	.165	.675
	TOI	-.024	.029	-.103	-.821	.415	-.082	.034

a. Dependent Variable: ANDI_ALL

Study decisions and accuracy

Descriptive Statistics

	Mean	Std. Deviation	N
C_w_Easy	.29366677	.203141093	65
C_w_Mod	.19062525	.207057522	65
C_w_Hard	.15540117	.135036719	65

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
difficulty	Sphericity Assumed	.396	2	.198	9.716	.000	.134
	Greenhouse-Geisser	.396	1.841	.215	9.716	.000	.134
	Huynh-Feldt	.396	1.924	.206	9.716	.000	.134
	Lower-bound	.396	1.000	.396	9.716	.003	.134
difficulty * TOI	Sphericity Assumed	.228	2	.114	5.591	.005	.082
	Greenhouse-Geisser	.228	1.841	.124	5.591	.006	.082
	Huynh-Feldt	.228	1.924	.119	5.591	.005	.082
	Lower-bound	.228	1.000	.228	5.591	.021	.082
Error(difficulty)	Sphericity Assumed	2.570	126	.020			
	Greenhouse-Geisser	2.570	115.975	.022			
	Huynh-Feldt	2.570	121.198	.021			
	Lower-bound	2.570	63.000	.041			

Correlations

		TOI	C_w_Easy
TOI	Pearson Correlation	1	-.260*
	Sig. (2-tailed)		.036
	N	65	65
C_w_Easy	Pearson Correlation	-.260*	1
	Sig. (2-tailed)	.036	
	N	65	65

*. Correlation is significant at the 0.05 level (2-tailed).

Correlations

		TOI	C_w_Mod
TOI	Pearson Correlation	1	.049
	Sig. (2-tailed)		.699
	N	65	65
C_w_Mod	Pearson Correlation	.049	1
	Sig. (2-tailed)	.699	
	N	65	65

Correlations

		TOI	C_w_Hard
TOI	Pearson Correlation	1	.202
	Sig. (2-tailed)		.106
	N	65	65
C_w_Hard	Pearson Correlation	.202	1
	Sig. (2-tailed)	.106	
	N	65	65

Study approaches

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	JOLs_Chosen	37.46773796	65	13.82347353	1.714590870
	JOLs_NotChosen	48.66310132	65	16.94320678	2.101546156
Pair 2	WrongJOL_Chosen	35.65916684	65	13.90486904	1.724686741
	WrongJOL_NotChosen	41.31935162	65	15.97635734	1.981623242

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	JOLs_Chosen - JOLs_NotChosen	-11.19536337	16.79923983	2.083689254	-15.35801118	-7.032715556	-5.373	64	.000
Pair 2	WrongJOL_Chosen - WrongJOL_NotChosen	-5.660184783	14.61332899	1.812560383	-9.281190410	-2.039179155	-3.123	64	.003

ANCOVA Choice* TOI

Descriptive Statistics

	Mean	Std. Deviation	N
JOLs_Chosen	37.46773796	13.82347353	65
JOLs_NotChosen	48.66310132	16.94320678	65

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Choice	Sphericity Assumed	80.045	1	80.045	.559	.457	.009
	Greenhouse-Geisser	80.045	1.000	80.045	.559	.457	.009
	Huynh-Feldt	80.045	1.000	80.045	.559	.457	.009
	Lower-bound	80.045	1.000	80.045	.559	.457	.009
Choice * TOI	Sphericity Assumed	13.718	1	13.718	.096	.758	.002
	Greenhouse-Geisser	13.718	1.000	13.718	.096	.758	.002
	Huynh-Feldt	13.718	1.000	13.718	.096	.758	.002
	Lower-bound	13.718	1.000	13.718	.096	.758	.002
Error(Choice)	Sphericity Assumed	9017.144	63	143.129			
	Greenhouse-Geisser	9017.144	63.000	143.129			
	Huynh-Feldt	9017.144	63.000	143.129			
	Lower-bound	9017.144	63.000	143.129			

Tests of Between-Subjects Effects

Measure: MEASURE_1 ...

	Type III Sum					Partial Eta
Intercept	3929.944	1	3929.944	23.028	.000	.268
TOI	34.230	1	34.230	.201	.656	.003
Error	10751.480	63	170.658			

Pairwise Comparisons

Measure: MEASURE_1

(I) Choice	(J) Choice	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
1	2	-11.195*	2.099	.000	-15.389	-7.002
2	1	11.195*	2.099	.000	7.002	15.389

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

JOLs (For items wrong on the initial test)

Descriptive Statistics

	Mean	Std. Deviation	N
WrongJOL_Chosen	35.65916684	13.90486904	65
WrongJOL_NotChosen	41.31935162	15.97635734	65

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Choice	Sphericity Assumed	89.130	1	89.130	.854	.359	.013
	Greenhouse-Geisser	89.130	1.000	89.130	.854	.359	.013
	Huynh-Feldt	89.130	1.000	89.130	.854	.359	.013
	Lower-bound	89.130	1.000	89.130	.854	.359	.013
Choice * TOI	Sphericity Assumed	259.730	1	259.730	2.489	.120	.038
	Greenhouse-Geisser	259.730	1.000	259.730	2.489	.120	.038
	Huynh-Feldt	259.730	1.000	259.730	2.489	.120	.038
	Lower-bound	259.730	1.000	259.730	2.489	.120	.038
Error(Choice)	Sphericity Assumed	6573.851	63	104.347			
	Greenhouse-Geisser	6573.851	63.000	104.347			
	Huynh-Feldt	6573.851	63.000	104.347			
	Lower-bound	6573.851	63.000	104.347			

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Choice	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Choice	Level 1 vs. Level 2	178.260	1	178.260	.854	.359	.013
Choice * TOI	Level 1 vs. Level 2	519.459	1	519.459	2.489	.120	.038
Error(Choice)	Level 1 vs. Level 2	13147.701	63	208.694			

Tests of Between-Subjects Effects

Measure: MEASURE_1 ...

	Type III Sum					Partial Eta
Intercept	2009.989	1	2009.989	11.878	.001	.159
TOI	277.151	1	277.151	1.638	.205	.025
Error	10660.919	63	169.221			

Pairwise Comparisons

Measure: MEASURE_1

(I) Choice	(J) Choice	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
1	2	-5.660*	1.792	.002	-9.241	-2.079
2	1	5.660*	1.792	.002	2.079	9.241

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Pillai's trace	.137	9.979[a]	1.000	63.000	.002	.137
Wilks' lambda	.863	9.979[a]	1.000	63.000	.002	.137
Hotelling's trace	.158	9.979[a]	1.000	63.000	.002	.137
Roy's largest root	.158	9.979[a]	1.000	63.000	.002	.137

Each F tests the multivariate effect of Choice. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

a. Exact statistic

Optimal choices

Descriptive Statistics

	Mean	Std. Deviation	N
HonJOL_Chosen	37.7870	15.09230	64
F_Recall_Hon	5.2500	3.51414	64

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.145[a]	.021	.005	15.05273

a. Predictors: (Constant), F_Recall_Hon

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	301.736	1	301.736	1.332	.253[b]
	Residual	14048.256	62	226.585		
	Total	14349.992	63			

a. Dependent Variable: HonJOL_Chosen

b. Predictors: (Constant), F_Recall_Hon

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	34.517	3.401		10.149	.000	27.719	41.316
	F_Recall_Hon	.623	.540	.145	1.154	.253	-.456	1.702

a. Dependent Variable: HonJOL_Chosen

Descriptive Statistics

	Mean	Std. Deviation	N
HonWrongJOL_Chosen	36.2146	15.27221	64
HonWrongFRecallChosen	1.5938	1.72488	64

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.155[a]	.024	.008	15.20982

a. Predictors: (Constant), HonWrongFRecallChosen

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	351.147	1	351.147	1.518	.223[b]
	Residual	14343.004	62	231.339		
	Total	14694.151	63			

a. Dependent Variable: HonWrongJOL_Chosen

b. Predictors: (Constant), HonWrongFRecallChosen

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B Lower Bound	95.0% Confidence Interval for B Upper Bound
1	(Constant)	34.033	2.598		13.100	.000	28.840	39.227
	HonWrongFRecallChosen	1.369	1.111	.155	1.232	.223	-.852	3.589

a. Dependent Variable: HonWrongJOL_Chosen

Honour vs. dishonour

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	F_Recall_Hon	5.2500	64	3.51414	.43927
	F_Recall_Dishon	5.2500	64	3.36650	.42081

Paired Samples Test

		Paired Differences Mean	Paired Differences Std. Deviation	Paired Differences Std. Error Mean	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper	t	df	Sig. (2-tailed)
Pair 1	F_Recall_Hon - F_Recall_Dishon	.00000	3.00264	.37533	-.75004	.75004	.000	63	1.000

Descriptive Statistics

	Mean	Std. Deviation	N
F_Recall_Hon	5.2500	3.51414	64
F_Recall_Dishon	5.2500	3.36650	64

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
honour	Sphericity Assumed	1.984	1	1.984	.436	.511	.007
	Greenhouse-Geisser	1.984	1.000	1.984	.436	.511	.007
	Huynh-Feldt	1.984	1.000	1.984	.436	.511	.007
	Lower-bound	1.984	1.000	1.984	.436	.511	.007
honour * TOI	Sphericity Assumed	2.065	1	2.065	.454	.503	.007
	Greenhouse-Geisser	2.065	1.000	2.065	.454	.503	.007
	Huynh-Feldt	2.065	1.000	2.065	.454	.503	.007
	Lower-bound	2.065	1.000	2.065	.454	.503	.007
Error(honour)	Sphericity Assumed	281.935	62	4.547			
	Greenhouse-Geisser	281.935	62.000	4.547			
	Huynh-Feldt	281.935	62.000	4.547			
	Lower-bound	281.935	62.000	4.547			

Tests of Between-Subjects Effects

Measure: MEASURE_1 ...

	Type III Sum					Partial Eta
Intercept	178.180	1	178.180	9.164	.004	.129
TOI	2.507	1	2.507	.129	.721	.002
Error	1205.493	62	19.443			